1. Self-Cognition in Transcendental Philosophy

Chong-Fuk Lau

1. Introduction

In the Preface to the First Edition of the *Critique of Pure Reason*, Kant explains his task as follows:

> reason should take on anew the most difficult of all its tasks, namely, that of self-knowledge [*Selbsterkenntnis*], and to institute a court of justice, by which reason may secure its rightful claims while dismissing all its groundless pretensions, and this not by mere decrees but according to its own eternal and unchangeable laws; and this court is none other than the *critique of pure reason* itself. (Axi-xii)

According to Kant, traditional metaphysics ended up in a battlefield of endless controversies because it tried to answer questions that transcend every capacity of reason. Kant, therefore, makes it his task to investigate the nature and limits of reason itself, attempting to institute a "court of justice" to settle metaphysical controversies once and for all. The *Critique* is a project of self-investigation and self-examination of reason. This self-reflexive character of Kant's project is also implied in the concept of *transcendentality*, referring primarily to a peculiar kind of *a priori* cognition: "not every *a priori* cognition must be called transcendental, but only that by means of which we cognize that and how certain representations (intuitions or concepts) are applied entirely *a priori*, or are possible" (A56/B80 f; cf. B25). In other words, transcendental cognition is the sort of *a priori* cognition that deals with the very possibility and conditions of cognition itself. Since these conditions are rooted in our faculty of cognition, "the word transcendental", for Kant, "never signifies a relation of our cognition to things, but only to the *faculty of cognition* [*Erkenntnisvermögen*]" (4:293).

Kant's *Critique* is a study of our own faculty of cognition and the principles derived therefrom. Kant offers a systematic division of the mind into different cognitive and non-cognitive faculties; among these the distinction between sensibility and understanding is most crucial to his epistemology. Sensibility and understanding are two stems of

human cognition that are both indispensable and mutually irreducible (A15/B29). While objects are given to us by means of sensibility, they are thought through the faculty of understanding. The former is subject to the forms of space and time, and the latter to the categories. These are the basic tenets of Kant's epistemology. The problem I would like to discuss is the theoretical status of these transcendental cognitions. Are they descriptions about the structure and operations of the human mind? Kant's *Critique* often gives the impression of describing a system of hidden psychological processes or cognitive mechanisms that constructs the world of appearances out of the manifold given from things in themselves. This psychological picture is seriously misguided. I will argue that Kant's discussion of the faculties of sensibility and understanding do not straightforwardly describe the human mind, but the investigation into the faculty of cognition is a conceptual analysis of the structure of cognition of the finite rational being as such.

2. Psychological Interpretation

Human beings are equipped with a highly sophisticated cognitive apparatus. We perceive the world through different senses and process the sense-data with an intelligent brain. Does Kant refer to the human cognitive system, when he speaks of the faculty of cognition? How do the faculties of sensibility and understanding differ from, say, the system of the five senses (sight, hearing, smell, taste and touch)? For Kant, the difference is fundamental. The latter refers to the cognitive apparatus of human beings that we know through empirical observation. Descriptions about the five senses are *a posteriori* and contingent. It is totally possible that human beings could have developed more or less than five senses, or different senses than the ones we now have. Nothing rules out the possibility that we someday may acquire the ability to perceive ultrasonic signals, so that we may be able to know what it is like to be a bat. Our possibility in perceiving the world is certainly constrained to a certain extent by the existing five senses, but they are not the *a priori* conditions that account for the possibility of objective cognition. Kant is well aware of the difference (A29/B44). The five senses do not belong properly to the subject of the *Critique*, but rather to that of the *Anthropology*.[1]

1 Kant discusses the five senses in some detail in *Anthropology from a Pragmatic Point of View* (7:153 f).

The faculties of sensibility and understanding as well as their spatiotemporal and categorial forms are the *a priori* conditions that the *Critique* has to deal with. Henry Allison calls them "epistemic conditions" in order to distinguish them from merely "psychological conditions."[2] Human beings are, for example, only capable of perceiving light of a certain wavelength. This certainly belongs to the conditions of human visual perception, but for Allison, "conditions of this sort are not epistemic in the relevant sense, because they have no objective validity or objectivating function."[3] In other words, these conditions are not indispensible for object formation. Objects appearing under a different color scheme or even without visual properties are conceivable, but the epistemic conditions of spatiotemporal sensibility and categorial understanding are so fundamental that no objects would ever be conceivable if human beings had not developed an appropriate cognitive system incorporating these features. The epistemic conditions, therefore, enjoy a privileged status compared to other psychological conditions.

However, why do space, time, and categories have the indispensible objectivating function, while others do not? The question cannot be answered unless the meaning of Kant's discussion about our faculty of cognition is correctly understood. The distinction between sensibility and understanding, together with their formal conditions, should not be understood as characteristics that, as Allison suggests, "reflect the structure and operations of the human mind."[4] Descriptions of sensibility and understanding are essentially different from all empirical descriptions of the human mind. They are not directly about any factual characteristics of a biological species that happens to have developed sophisticated cognitive abilities. All factual descriptions of the human mind are bound to be contingent and known empirically. The conditions of sensibility and understanding themselves cannot be identified through empirical knowledge, including discovery by psychological or introspective observation; otherwise, they would not be *a priori* conditions of experience and objectivity. The human mind can certainly be studied as an empirical cognitive apparatus, but this belongs to the task of cognitive science or psychology and not to Kant's transcendental philosophy.

2 H. E. Allison, *Kant's Transcendental Idealism: An Interpretation and Defense*, second edition (New Haven: Yale University Press, 2004), 11.

3 Allison, *Kant's Transcendental Idealism*, 12.

4 Allison, *Kant's Transcendental Idealism*, 11.

Although Kant does take the "faculty of cognition" to be the "object" of transcendental cognition, it cannot be understood as a cognitive apparatus that functions in a specific way to produce knowledge. The psychological interpretation not only mistakes the faculty of cognition for an empirical object, but more importantly fails to acknowledge that the *conditions* of the possibility of empirical reality themselves cannot be part of empirical reality. The search for the "structure and operations of the human mind" is misguided, if it aims to reveal the empirical workings of the human mind. Every process or operation of the human mind operates and takes place *in time*. Now, if temporality belongs to *a priori* formal conditions of appearance, then the formal conditions that account for the possibility of appearance cannot themselves be temporal processes, nor anything determinable in time. In Kant's words, "the subject, in which the representation of time originally has its ground, cannot thereby determine its own existence in time" (B422). It does not make sense to explain why and how things must be ordered in a temporal series by describing processes that themselves take place in time. A temporal process cannot be a condition of temporality. The "faculty of cognition" cannot stand for a system of cognitive processes or operations that takes place in the human mind (or the brain) to organize formless data into spatiotemporally and categorially structured appearance. This psychological interpretation mistakes Kant's project for "a certain *physiology* of the human understanding (by the famous Locke)" (Aix).

3. The Nature of Epistemic Conditions

If the epistemic conditions cannot be taken as describing the empirical mind or the phenomenal self, does this mean they refer to something that lies behind the empirical realm of appearance? Are the sensibility-understanding distinction and their formal conditions properties of the noumenal self? Such an interpretation would be in direct conflict with Kant's criticism of traditional rational psychology's attempt to prove the substantiality, simplicity, unity and independent existence of the thinking subject from the sole text of "I think." Kant's analysis in the Paralogisms unveils the illusion of this metaphysical doctrine, explaining why categories such as substance and causality cannot be legitimately applied to the thinking subject. Kant's basic epistemological principle rules out the possibility of knowing anything about things in

themselves or noumena. The principle of noumenal ignorance applies to the same extent to the thinking subject itself as to any other external things. It is illegitimate to ascribe the faculties of sensibility and understanding to the noumenal self, since this would presuppose the noumenal self as a substance, with properties attributable to it. It would also make no sense to talk about "operations" or "constructions" of the noumenal self, because temporal and causal categories cannot be applied to it.

According to Kant, no rational self-cognition is possible. I can become an object of cognition to myself only through empirical means such as inner sense. "Yet", as Kant once conceded, "the human being, who is otherwise acquainted with the whole of nature solely through sense, knows [*erkennt*] himself also through pure apperception" (A546/B574). Using the word "know" or "cognize" [*erkennen*] in this context seems hardly reconcilable with Kant's basic epistemological tenets. Kant may have chosen a wrong word here, but the concept of "pure apperception" points to another dimension of the problem. Kant does allow a kind of non-empirical access to the faculty of cognition (i. e., a form of intellectual self-consciousness). Although understanding is not a faculty of intuition, it can be "conscious" of the unity of its action "even without sensibility" (B153). The possibility of non-empirical self-consciousness is crucial to Kant's epistemology. It concerns the pure, original, and transcendental apperception in contrast to the empirical self-consciousness that is in constant flux (A107). Pure apperception refers to "the representation *I think*, which must be able to accompany all others [representations] and which in all consciousness is one and the same, cannot be accompanied by any further representation" (B132). This non-empirical self-consciousness serves as a unifying condition for all representations that can be attributed to a thinking subject.

In Kant's epistemology, knowledge requires both concept and intuition. As "the consciousness of myself in the representation *I* is no intuition at all, but a merely intellectual representation of the self-activity of a thinking subject" (B278), self-consciousness is still far from being self-cognition (B158). In order to acquire self-cognition, self-intuition is required; this, however, can only be given empirically. Therefore, as argued in the Paralogisms, no rational cognition whatsoever about the thinking subject can be derived from pure apperception. But Kant does admit the possibility of establishing *a priori* cognitions from pure apperception; he calls it "the *transcendental* unity of self-consciousness in order to designate the possibility of *a priori* cognition from it"

(B132). A passage from the *Anthropology* (7:135) further explains the role of pure apperception in logical cognition:

> If we consciously represent two acts: inner activity (spontaneity), by means of which a *concept* (a thought) becomes possible, or *reflection*; and receptiveness (receptivity), by means of which a *perception* (*perceptio*), i.e., empirical *intuition*, becomes possible, or apprehension; then consciousness of oneself (*apperceptio*) can be divided into that of reflection and that of *apprehension*. The first is a consciousness of understanding, *pure* apperception; the second a consciousness of inner sense, *empirical* apperception. In this case, the former is falsely named *inner* sense.—In psychology we investigate ourselves according to our ideas of inner sense; in logic, according to what intellectual consciousness suggests.

In order to acquire knowledge of oneself, including one's internal mental state, one has to rely on what is given empirically in inner (and outer) sense. I cannot know anything about myself through pure apperception alone, yet Kant suggests this kind of intellectual consciousness is the source of logical knowledge.

Although Kant's terminology sometimes gives the impression of being psychological, a comparison of his transcendental logic with the general logic will show that his discussion of sensibility and understanding does not refer to mental structures or operations. For Kant, logic is the "science of the necessary laws of the understanding and of reason in general" (9:13; cf. A53/B77). This certainly does not mean pure logic studies how understanding and reason in fact think or operate. There is a branch of logic that "is directed to the rules of the use of the understanding under the subjective empirical conditions that psychology teaches us" (A53/B77), but it is called *applied* logic. Abstracted from all psychological and empirical elements, pure logic is concerned with how understanding and reason *ought to* work (i.e., with the necessary *normative* principles for thought). Logical principles are necessary not in the sense that we cannot think otherwise; instead, their necessity carries normative bindingness (*Verbindlichkeit*). What Kant calls *transcendental logic* also "has to do merely with the laws of the understanding and reason, but solely insofar as they are related to objects *a priori*" (A57/B81 f). Just as in the case of pure general logic, transcendental logic does not describe how understanding and reason work, but *prescribes* how their concepts or ideas *ought to* be applied in order to form objective cognition. That is why Kant emphasizes that the transcendental deduction of categories is not primarily concerned with factual questions

(*quid facti*), but rather with questions about entitlement or legitimacy of the use of categories (*quid juris*) (A84/B116).[5]

4. Self-Cognition as Cognition of Cognition

Pure apperception is the consciousness of understanding or spontaneity that enables concepts to be thought. What is revealed in pure apperception is not any internal mental state, but the content of conceptual relations: they are pure and non-empirical because they are not determined in a temporal-causal series. The conceptual-normative principles determined by the spatiotemporal and categorial forms of sensibility and understanding, respectively, can be ascribed to the faculty of cognition, but the latter refers to neither the phenomenal nor the noumenal self. It can be said to belong to the transcendental self or subject, and is

> nothing but the simple and in content for itself wholly empty representation *I*, of which one cannot even say that it is a concept, but a mere consciousness that accompanies every concept. Through this I, or He, or It (the thing), which thinks, nothing further is represented than a transcendental subject of thoughts = x (A345 f/B404).

The transcendental subject is not a thinking human being that exists in time and processes mental representations with the brain, but only a *theoretical construct* that accounts for the unity and legitimacy of epistemic conditions. The transcendental apperception "I think" is not a factual description of my thinking activity either, but merely a "logical function" (B143, B428). I can certainly use the statement "I think" to describe the empirical fact that I am now thinking, "but the proposition 'I think,' insofar as it says only that *I exist thinking*, is not a merely logical function, but rather determines the subject (which is then at the same time an object) in regard to existence, and this cannot take place without inner sense" (B429). The transcendental subject is thus not the I that exists in the empirical world and engages in the activity of thinking. It is the ground that makes the indispensible spatiotemporal and categorial structure of objects possible.

---

5 Accordingly, the distinction between understanding and reason in the Transcendental Logic does not primarily address two different cognitive components that human beings in fact possess, but distinguishes between two *types* of normative principles that are valid for making *judgments* and *inferences*, respectively.

Accordingly, the faculties of sensibility and understanding should not be understood primarily as the cognitive components for receiving and processing sense impressions. While understanding provides the normative principles for the application of concepts in judgments, sensibility is responsible for the normative principles that govern the reference to objects. Sensibility and understanding represent two basic and mutually irreducible *functions* for the cognition of objects (i.e., *reference* and *description*).[6] Understanding does not refer to objects directly, but thinks them by applying *concepts* that are nothing but "predicates of possible judgments" (A69/B94), whereas sensibility is responsible for nonconceptual, "immediate" reference to individual objects that predicates can be ascribed to. Jaakko Hintikka suggests "that Kant's notion of intuition is not very far from what we would call a singular term."[7] Similarly, Wilfred Sellars understands Kant's intuition as a function of representing individuals or "*thises*."[8]

In this connection, space and time are said to be the *a priori* forms of sensibility, not because sensibility receives and processes sense impressions through a spatiotemporal spectacle, but because reference to individual objects has to rely on spatiotemporal coordination. Sellars puts the point as follows:

> To intuit is to represent a *this* … Space and time are 'forms of intuition,' not by virtue of being attributes of or relations between things or events in nature, but by virtue of the fact that the logical powers distinctive of 'this' representings are specified in terms of concepts pertaining to relative location in space and time.[9]

For Kant, the spatiotemporal framework underlies the possibility of individuation.[10] Space and time form a normative framework that enable

6 This is similar to the *demonstrative* and *descriptive conventions* that J. L. Austin takes to be necessary for making meaningful statements. See J. L. Austin, "Truth," in *Philosophical Papers*, eds. J. O. Urmson and G. J. Warnock (Oxford: Clarendon Press, 1961), 121 f.

7 J. Hintikka, "On Kant's Notion of Intuition (Anschauung)", in *The First Critique: Reflections on Kant's Critique of Pure Reason*, ed. T. Penelhum and J. J. Maclntosh (Belmont, Ca.: Wadsworth Publishing Company, 1969), 43.

8 W. Sellars, "Sensibility and Understanding", in *Science and Metaphysics: Variations on Kantian Themes* (London: Routledge & Kegan Paul, 1968), 3.

9 Sellars, "Some Remarks on Kant's Theory of Experience", in *Essays in Philosophy and Its History* (Dordrecht: D. Reidel Publishing Company, 1974), 54.

10 Kant does not subscribe to Leibniz's principle of the identity of indiscernibles. Two drops of water without any inner difference, for instance, can still be held

identifiable references to individual objects to be made. Accordingly, the necessary cooperation of sensibility and understanding in Kant's epistemology can be pinned down to the idea that the possibility of objective cognition presupposes two basic types of normative principles that govern our references to and descriptions of objects, respectively. The point of attributing these forms to the transcendental subject is to express the theoretical status of these forms—that is, their essentially *normative* character. Normative forms are not given from things outside, but are something we *prescribe* to the empirical world. In this sense, "objects must conform to our cognition" (Bxvi). The crucial idea in Kant's Copernican revolution is therefore, as Robert Brandom suggests, "a normative turn."[11]

Kant's aim is to account for the possibility of objective cognition and reality; this presupposes the epistemic, normative principles discussed previously. If human beings are *in fact* capable of acquiring objective knowledge, our cognitive apparatus must function in a way that complies with the norms. Then, our cognitive apparatus must have sufficiently incorporated the epistemic conditions into its working mechanisms. But whether and to what extent this is the case is a matter-of-fact question that can only be decided by empirical (psychological or cognitive) research. Similarly, whether I am making a valid inference or whether I am thinking correctly according to the laws of logic is a factual matter, while the validity of the logical laws itself is of a normative nature. Insofar as we have reason to believe that human beings do have objective knowledge, we are entitled to assume our cognitive apparatus can be adequately described by the spatiotemporal and categorial forms of sensibility and understanding. But Kant's epistemology is not confined to the human cognitive system. It does not depend on the particular biological, psychological, or cognitive structures of Homo sapiens. Kant's project is a conceptual analysis that should be valid to every being that is capable of acquiring objective knowledge, or at least valid to *all finite rational beings*. Kant is relatively conservative at this point, but he does say that "it is also not necessary for us to limit the kind of intuition in space and time to the sensibility of human beings; it may well be that all finite thinking beings must necessarily agree with human be-

to be numerically different by virtue of their different spatiotemporal locations (A263/B319 f).

11 R. B. Brandom, *Tales of the Mighty Dead: Historical Essays in the Metaphysics of Intentionality* (Cambridge, Ma.: Harvard University Press, 2002), 21.

ings in this regard (though we cannot decide this)" (B72). In regard to the condition of the validity of reason's ideas, Kant even claims that "given the nature of our (human) cognitive faculty or even the concept that *we can form* of the capacity of a finite rational being in general, we cannot and must not conceive otherwise" (5:401). Kant is concerned with the universal normative conditions, but his analysis seems to have taken a "shortcut", as if he were straightforwardly describing the structure and operations of the human mind. This is a major reason Kant's theory often gives the impression of being psychological. If Kant's investigation into the faculty of cognition provides a sort of *self*-cognition, then it is not cognition of the self, but rather cognition of cognition.